

COMPARISONS WITH PREVIOUS ESTIMATES AND IMPLICATIONS FOR REVISIONS

INTRODUCTION

The 2001 Population Report provides population data for Census Day 2001 and 30 June 2001. It also shows the previously published figures for mid-2000 and a provisionally revised set of mid-2000 figures that are consistent with the 2001 data.

The revised figures are some 56,000 lower than the previously published figures. This paper explains the extent of the differences by age and sex and gives some guidance on the implications for revising figures for earlier years in line with the 2001 Census data.

A key implication is the need to revise the mid-1991 figures since there is now evidence that these were too high. Consequently there is likely to be a need to revise the figures for each year since 1982.

RESULTS

Census Day

The 2001 Census provides the best estimate of the population in 2001. There is a difference of some 50,000 between Scotland's Census Day population count and the previously published 2000 mid-year estimate. The difference is primarily attributable to males, particularly in men aged 20-39.

As a result, the ratio of men to women is lower in these age groups than had been previously estimated. The next section provides a wider analysis of changes by age, sex and area.

2000 mid-year estimates and "error"

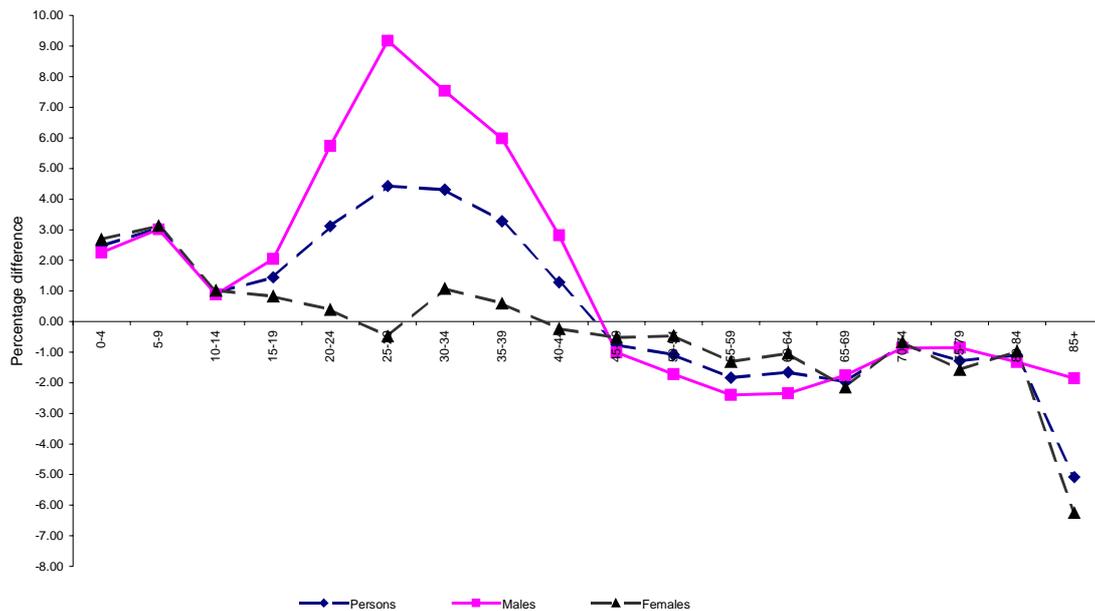
A set of provisionally revised 2000 mid-year estimates has been produced to provide a comparison with the previously published 2000 figures to give a measure of the "error" in the previous estimate as revealed by the Census. They have also been produced to provide a valid comparison between 2000 and 2001 figures, but this section will focus on exploring the "error" in the previous 2000 estimates and possible reasons for it.

These provisionally revised estimates were calculated by 'rolling back' the 2001 mid-year estimates, i.e. the 2001 population was aged back by a single year of age at each age and sex, deaths in the previous 12 months were added, births were subtracted and an estimate of migration was subtracted to arrive at a provisional revised set of 2000 mid-year estimates. These are considered provisional revisions as further revisions to the 2000 estimates may be necessary in the light of the 2001 Census figures.

Age/sex "error"

Table 1 and **Figure 1** compare the previously published 2000 figures with the provisionally revised 2000 figures by age and sex. Although males account for all of the difference an analysis of the differences across age groups shows some interesting changes.

Figure 1 Percentage error in original 2000 estimates

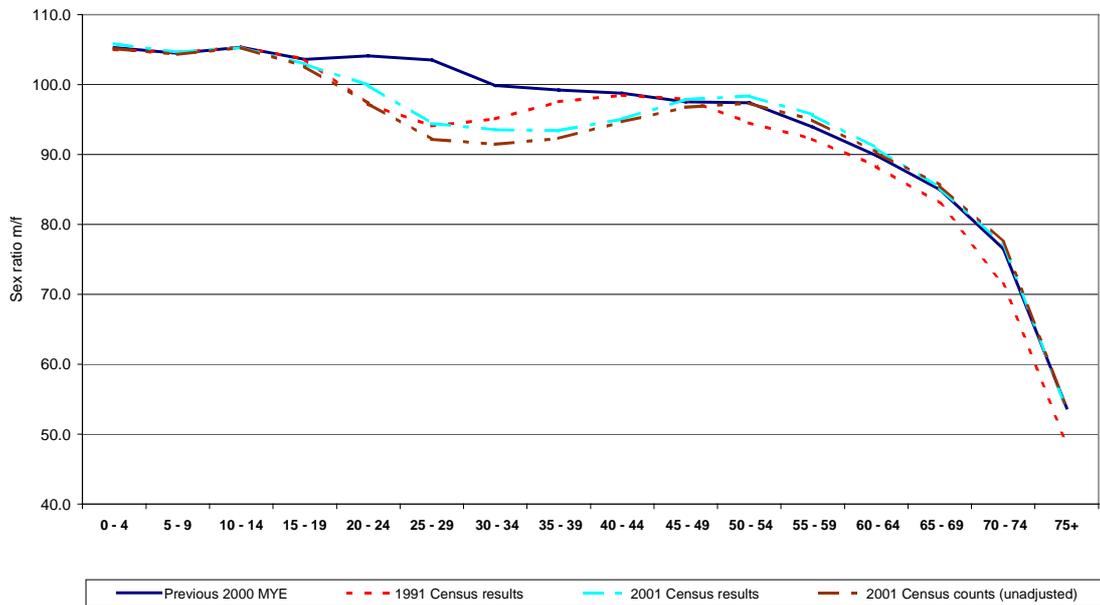


The main differences are in the population estimates for men, particularly at ages 20-39. This difference occurs because the previously published estimates have overestimated the males aged 20-39 by 53,000 or 7 per cent. There are differences throughout female age groups, but on the whole these differences are small except for women aged over 85 where previous estimates were too high. The ratio of men to women is therefore lower at ages 20-39 than had been estimated in the past. Women now outnumber men from the age of 22 compared with age 34 in the previously published estimates.

There are also significant differences between the two estimates for males aged over 45. These are the results of underestimating, rather than overestimating, the population in the previously published estimates. Although these differences are smaller than the differences in males aged 20-39, they are, on the whole, larger than for females.

The pattern of sex ratios found in the 2001 Census is very similar to those found in 1991 Census data before adjustments for under-enumeration were included in preparing the 1991 mid-year estimates. These ratios, along with the 2001 Census ratios after adjusting for underenumeration and the previously published 2000 ratios are shown in **Figure 2**. At the time of the 1991 Census these findings were discounted given the problems with the post-enumeration survey and it was thought that there was no credible evidence to support the observed pattern. Alternative sex ratios, based on information from previous Censuses, were adopted in the mid-1991 population estimates. (for more details, see 1991 CVS coverage report ISBN 0 11 691591 9).

Figure 2: Comparison of sex ratios



Given the consistency between the findings in the 1991 and 2001 Census and the lack of evidence from the much more comprehensive Coverage Survey conducted in 2001 about missing young males, there is good reason to believe that there really was fewer young men than women resident in Scotland in 1991. The adjustments made to the 1991 Census data were therefore too high. Consequently there will be a need to revise the 1991 counts and probably to revise the data for all years from 1982 to 2000. More information on revisions is given in the Further Work section.

Council area “error”

The numbers of councils falling within a range of errors is shown in **Figure 3**. This shows that for 13 out of 32 councils the “error” in the previously published estimate is small - less than one per cent - and a further 12 councils have an error of less than 2 per cent.

Figure 3 Degree of change for council areas between previously published and provisionally revised mid-2000 population estimates

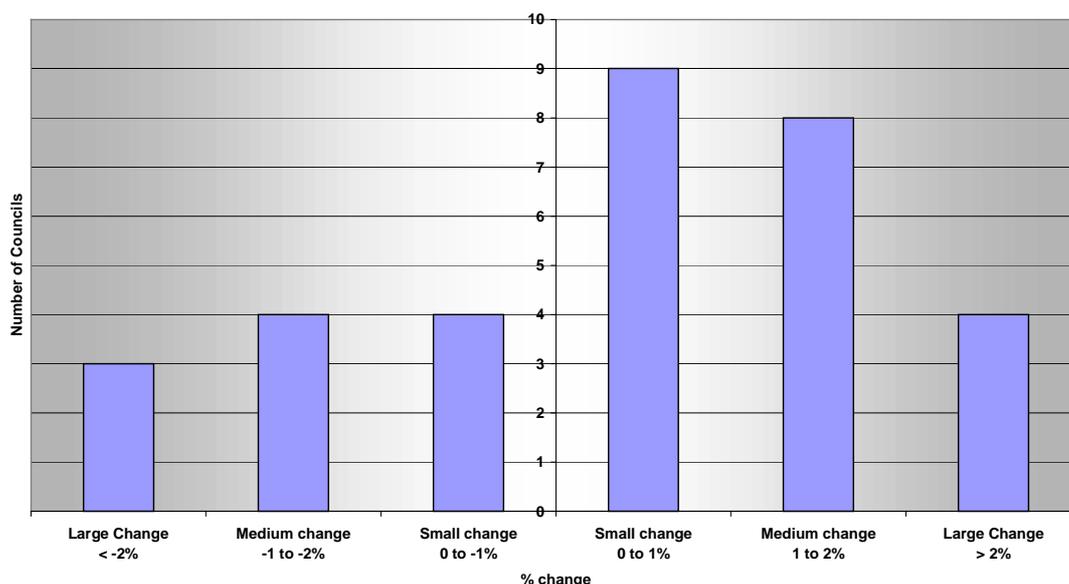


Table 2 shows the differences between the two 2000 population estimates by council area. The councils with the largest overestimate in the previously published 2000 figures are Glasgow (5.9%) 36,000 high, Renfrewshire (2.1%) 3,700 high and North Lanarkshire (2.1%) 6,800 high. Councils with the greatest underestimate in the previously published 2000 figures are Dundee City (3.0%) 4,300 low, Argyll & Bute (2.9%) 2,600 low and Moray (2.3%) 1,900 low.

EXPLAINING THE DIFFERENCE

In Scotland, and the rest of the UK, the mid-year population estimates are based on an accepted and widely employed demographic method, the cohort component method. Using this method, estimates are produced by updating from a census base allowing for the births, deaths and net migration that has occurred since the census. The mid-year estimates are the best estimates that can be made annually on a nationally consistent basis but inevitably they are subject to error which can accumulate between benchmarks. This section explores the two sources of “error” in the mid-year estimates in more detail and examines their respective contributions to the difference of some 56,000 between the previously published and the provisionally revised 2000 population estimate.

There are two main components of population change – natural change and migration. Natural change is the excess of births over deaths and this is estimated using data from the civil registration system - the compulsory registration of births and deaths. This data is considered to be virtually complete and therefore easy to estimate, in contrast to the estimation of migration. There is no comprehensive registration of migration in the UK, either of moves to or from the rest of the world, or of moves within the UK. Thus estimates of migration have to be based on survey data and the best proxy data that exist.

Any errors in estimating annual population change due to migration will remain in the estimates and will be added to the errors in subsequent mid-years. Thus, the mid-2000 population estimates will contain more error due to migration measurement in the 1990s than earlier mid-year estimates during the 90s, which incorporate fewer years of estimated migration change.

Another source of error in the mid-2000 estimates is the adjustments that were made to the 1991 Census base for underenumeration in the 1991 Census. The results of the 2001 Census provide an additional data source from which to assess the plausibility of underenumeration adjustments made in 1991.

This section explores these two sources of error in more detail and examines their contribution to the difference of some 56,000.

Migration

It is expected that the difference of some 56,000 between the previously published and the provisionally revised 2000 mid-year estimates is largely attributable to errors in previous estimates of migration in the eighties and nineties. It has been well documented that migration is the most difficult component of population change to estimate because of the lack of a wholly reliable source of data on migration. Estimates of net migration are the difference between large flows of migrants into and out of Scotland based on survey data (International Passenger Survey), or, for movements within the UK, from available administrative data (National Health Service Central Register recording the moves of patients throughout the UK). These estimates of net migration will, therefore, be very sensitive to measurement error in the flows in each direction and even more sensitive for particular age/sex groups.

For Scotland there are two key flows of migrants: migrants moving within the UK; and migrants to and from the rest of the world. Initial investigations at the UK level suggest that moves to and from the rest of the world (international migration) are the primary area of concern. This is also likely to be true for Scotland. Therefore this section focuses largely on international migration.

International migration

International migration is difficult to measure both for in flows and out flows. Estimates are derived from the International Passenger Survey (IPS) (see here for more information on the IPS) The IPS is the richest source of information on international migration into and out of the UK and areas within the UK. It is a continuous voluntary sample survey conducted by the Office for National Statistics (ONS), based on face-to-face interviews, which covers the principal air, sea and channel tunnel routes between the UK and countries outside the British Isles.

It is a well designed survey that provides valuable data but as well as migration, it is also designed to measure tourism and the contribution of travel expenditure to the balance of payments and, therefore, the sample of migrants contacted is small, particularly for Scotland. Since the figures obtained are based on a small number of contacts, the estimates are subject to larger sampling and non-sampling errors and are considered less reliable.

Another potential source of over or underestimation in the IPS is that it measures intentions which may not be realised. A person who comes here stating an intention to

migrate who then subsequently leaves after only a short stay would not be counted out as a migrant. On the other hand, the opposite case where people state their intention is to visit and then stay longer thus becoming a migrant to Scotland, is at least partly allowed for by the addition of Home Office data in respect of visitor switchers. Another difficulty arises where people leaving state that their intention is to be away for a short time but then they do not return and no adjustment is made for such cases as happens with the equivalent migrants moving to Scotland.

It is recognised that statistics on immigrants are difficult to compile and quality assure with confidence, but there is some limited evidence from registrations with NHS General Practitioners which indicates that immigration has not contributed significantly to the “error” in estimating migration.

There is less information about outward migration and, in the past, we have used the IPS on its own to produce complete estimates of emigration. The results from the 2001 census suggest that this approach may have led to an underestimation of out migration.

There is evidence to suggest that the international migration data sources have not been capturing all out migration. Data from a selection of key recipient countries which include Australia, Canada, New Zealand and the USA show that there are significant numbers of UK born people among immigrants to those countries. Most of these emigrants from the UK are young to middle aged adult men. Due to definitional differences and restrictions on the data that are available, it is not easy to carry out comparisons of the data from these countries with the estimates that have been used in the UK let alone for Scotland. However, one country where there is some definitional similarity, making comparisons possible, is Australia. Data from the Australian immigration control system shows that for the period from 1981 onwards there has been a net flow of approximately 260,000 permanent and long-term settlers to Australia of people born in the UK. Our combined data sources estimate 46,000 less than this. In addition, there has been a large increase in short-term arrivals in Australia by people born in the UK aged 25-44; these people are also more likely to be men than women. Some of these short-term visitors may stay on. Thus there is a strong likelihood that the underestimate of migration from UK to Australia exceeds the 46,000 that can be measured from the permanent and long-term flows.

In short, then, evidence does exist from recipient countries to support the proposition that there is out migration from UK, particularly of young to middle-aged adult males, that is not being captured by the data sources that have in the past been used to estimate migration. Similar analyses are not readily possible for other countries. It may well be that what is seen for Australia will also be seen for other countries.

There is little evidence from recipient countries about migration flows from Scotland rather than from UK.

Within UK migration

Migration between health boards within the UK is measured by the National Health Service Central Register (NHSCR), which is notified when a patient transfers to a new NHS doctor in a different health board area. Counts of these re-registrations are used as proxy indicators for moves within the UK and account for about 70 per cent of all migration movements into and out of Scotland. It is thought that the NHSCR re-registration data provide good indicators of inter-censal trends in migration but there

are certain characteristics of the NHSCR and the way it is used that may account for some of the 56,000 “error” at the Scotland level.

In particular the time taken between moving and re-registration which varies by sex and age-group. For example, young children and their mothers and the elderly usually re-register quite quickly after moving, while young adult males usually are less likely to re-register quickly than females of the same age. It is likely that the lag of re-registering for young adult males will have an impact on flows between Scotland and the rest of the UK, but it is thought that the impact of these lags may have been balanced out by direction and over time. However, further analysis of these flows will be undertaken to assess the impact and recommend improvements in the use of the re-registration data.

The mid-1991 to mid-2000 population estimates

The errors in estimating migration in the 1980s will also have affected the adjustments made to 1991 Census counts to take account of underenumeration. The adjustments to 1991 Census counts were incorrect and added some people, particularly young men, to the population whom we now believe had actually migrated out of Scotland rather than having been missed by the 1991 Census. This section looks at the possible errors in the estimates of underenumeration made for the 1991 Census.

In usual circumstances, it would be expected that the estimates for mid-1991 to mid-2000 would be based on the 1991 Census together with allowance for underenumeration in that census. However, there were difficulties with the assessment of underenumeration in the 1991 Census. The 1991 post enumeration survey (the Census Validation Survey, CVS) was a combined quality and coverage survey in which about 6,000 households across Great Britain were selected for the sample. It was not completely independent of the census and it was carried out to the same basic procedure as the original census. Therefore, any limitations inherent in the census would remain in the CVS. In particular, there may have been some people who were not counted by either the census or the CVS.

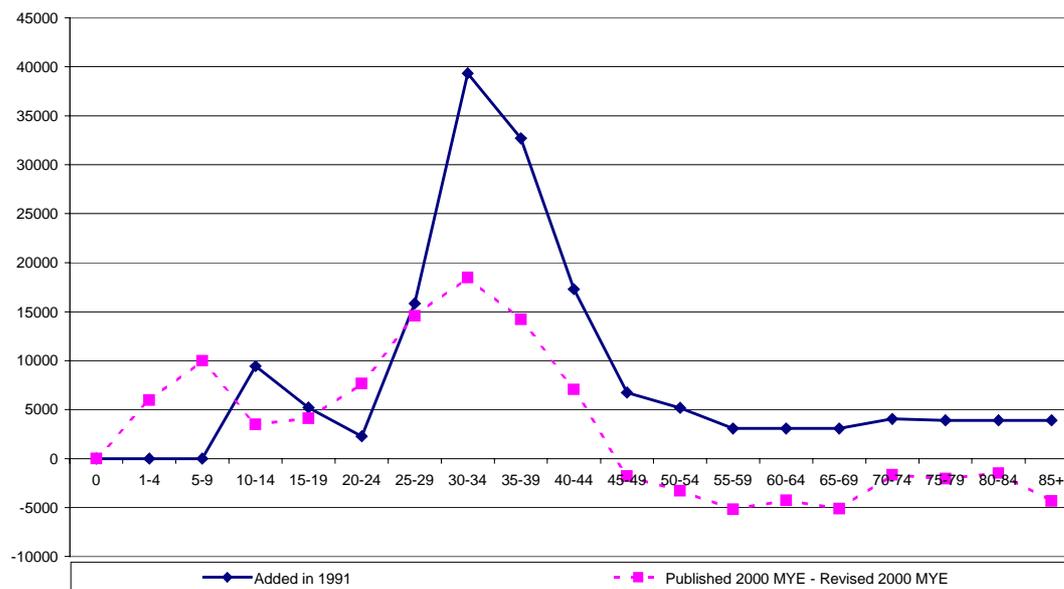
Given the problems with the 1991 CVS and the small sample size in Scotland, it was decided that the total figure for the national population of Scotland in mid-91 would be a number that was very close to that derived by updating from the 1981 Census. This was judged to be more plausible than the national total based on the results of the 1991 census. It was only this one figure that was used - the national total of the rolled-forward from 1981 estimates. The 1991 Census results were used to help derive the area distribution and the age and sex profile to ensure that the mid-91 estimates by age, sex and area agreed to the national total.

Thus the census and CVS results had to be adjusted for underenumeration that had not been identified by the CVS. To achieve this Scotland was divided into 5 areas and differential adjustments for underenumeration were made by area, sex and age group. Certain areas, namely Glasgow City, were given quite large adjustments, particularly to men aged 20-29.

Table 3 compares the levels of underenumeration adjustments by age and sex for 1991 and 2001. What is striking is that some of the largest differences between the 1991 and

2001 Census adjustments correspond to the age/sex area differences between the previously published and provisionally revised 2000 estimates of population. This suggests that the mid-1991 adjustments for underenumeration and, by implication, the national total updated from 1981, were not, in fact, correct. That in itself is not surprising. With hindsight it is perhaps to be expected that the national total produced by updating from 1981 would be subject to some error as a consequence of estimating migration for the 10 years during the 1980s.

Figure 4 Numbers added in 1991 (aged 10 years) compared to difference between previously published and provisionally revised mid-2000 population estimates



Based on the evidence now available, it is estimated that the adjustments made to the mid-1991 estimates and, accordingly, all population estimates produced during the 1990's, have been too high. In view of these findings, the population estimates for mid-1991 will be reconstructed, taking this discrepancy into account. The methodology for doing this will be presented to PAMS (Population and Migration Statistics) Group of the SCOTSTAT consultative arrangements and will be available on the GROS website.

Reconstructing the mid-1991 estimates means that it will also be necessary to revise the population estimates for the 1980's as well as the 1990's. It is likely that this work will be carried out to ensure that the revised set of population estimates for 1982-2000 will be available in February 2003.

The concept of usual residence

The 2001 census results are based on the One Number Census methodology incorporating full allowance for all those missed by the census. They provide the best possible population base for subsequent estimation of the resident population. A traditional census on its own without the use Census coverage Survey would have produced far less robust estimates in 2001 of those usually resident in an area. The issues we have had to grapple with in conducting the census have shown that there are increasingly complex issues in defining the concepts required for understanding the

population and for measuring those concepts robustly. There have been big changes in society in the last 10-20 years and one way in which this is manifesting itself is in where people consider themselves to live. For instance, some people have become more mobile since 1991, travelling around the world and staying temporarily in different places. Results becoming available from the 2001 census provide an exciting opportunity to learn something about these societal changes and their implications for population statistics. Our future development programme will take this opportunity in order to ensure that our system of population statistics in the future remains robust and relevant.

FURTHER WORK

Revised series of population estimates

A detailed study of the data used in the construction of the population estimates series for the 1990s and 1980s will now be carried out. This work will be an important input to the preparation of a revised population series for Scotland for the years 1982 - 2000 to provide a consistent time series between the 1981 and 2001 figures. Additionally, these analyses may suggest methodological improvements that can be made in preparing estimates of migration which would be incorporated in future estimates of the population. A revised set of population estimates for 1982-2000 will be available in February 2003.

Improving migration

The lack of a reliable source of data from which to estimate international migration is a particular concern, and a UK National Statistics project has been initiated to review the quality of international migration statistics. This project will review the data and methods with the aim of recommending improved approaches to measuring these migration flows, and is scheduled to report in 2003.

In addition, further work will be undertaken to examine migration between England & Wales and Scotland with the aim of improving the estimates of migration between the countries, particularly for young men where it is thought to be most deficient. This will be partially covered by work reconstructing the population series for the eighties and nineties and it is hoped that some improvements can be implemented before the 2002 mid-year estimates to be published in April 2003.

As well as the work identified above on Scotland level migration, work continues on a new method for estimating migration for council areas based on information from the Community Health Index (CHI). This work will be completed by the end of 2002 and a paper summarising the new method will be available on the GROS website.

Population projections

A set of 2002-based projections, which take account of the 2001 Census, is planned for publication in November 2003. The 2000-based population projections for Scotland and its administrative areas, published by the Registrar General earlier this year, used the previously published 2000 mid-year estimates as a base, which are some 50,000 higher than the provisionally revised 2000 mid-year estimates published in **Table A1**. While the 2000-based projections are in need of updating, the general picture of a slowly declining population is still relevant. To obtain revised projections which take

account of the results from the 2001 Census prior to the availability of the 2002-based projections, it may be best to apply the previously projected percentage changes to the provisionally revised base.

Table 1 Comparison of previously published and provisionally revised 2000 MYE

	Previously published 2000 MYE			Provisionally revised 2000 MYE			Error in previously published		
	Persons	Males	Females	Persons	Males	Females	Persons	Males	Females
All ages	5,114,600	2,484,687	2,629,913	5,058,200	2,428,100	2,630,100	1.10	2.28	-0.01
0-4	288,310	147,890	140,420	281,206	144,553	136,653	2.46	2.26	2.68
5-9	322,075	164,575	157,500	312,199	159,620	152,579	3.07	3.01	3.12
10-14	325,762	167,108	158,654	322,666	165,619	157,047	0.95	0.89	1.01
15-19	322,190	163,940	158,250	317,541	160,593	156,948	1.44	2.04	0.82
20-24	318,895	162,651	156,244	308,957	153,314	155,643	3.12	5.74	0.38
25-29	343,750	174,842	168,908	328,522	158,799	169,723	4.43	9.18	-0.48
30-34	402,811	201,258	201,553	385,471	186,089	199,382	4.30	7.54	1.08
35-39	416,073	207,223	208,850	402,432	194,821	207,611	3.28	5.98	0.59
40-44	376,112	186,891	189,221	371,300	181,629	189,671	1.28	2.82	-0.24
45-49	331,263	163,577	167,686	333,792	165,235	168,557	-0.76	-1.01	-0.52
50-54	342,588	169,042	173,546	346,303	171,948	174,355	-1.08	-1.72	-0.47
55-59	278,346	134,820	143,526	283,459	138,049	145,410	-1.84	-2.40	-1.31
60-64	259,361	122,744	136,617	263,669	125,630	138,039	-1.66	-2.35	-1.04
65-69	234,484	107,709	126,775	239,101	109,598	129,503	-1.97	-1.75	-2.15
70-74	205,168	88,961	116,207	206,715	89,726	116,989	-0.75	-0.86	-0.67
75-79	164,270	65,319	98,951	166,383	65,877	100,506	-1.29	-0.85	-1.57
80-84	99,467	33,896	65,571	100,555	34,346	66,209	-1.09	-1.33	-0.97
85+	83,675	22,241	61,434	87,929	22,654	65,275	-5.08	-1.86	-6.25

Table 2 Comparison of previously published and provisionally revised 2000 mid-year estimates

Area	Population estimates: 30 June 2000		Percentage error in previous 2000 Estimate	Estimated Population 30 June 2001
	previously ¹ published	provisionally revised		
SCOTLAND	5,114,600	5,058,200	1.1	5,064,200
Aberdeen City	211,250	213,470	-1.1	211,910
Aberdeenshire	227,200	227,030	0.1	226,940
Angus	109,180	108,800	0.3	108,370
Argyll & Bute	88,790	91,390	-2.9	91,300
Clackmannanshire	48,460	48,160	0.6	48,070
Dumfries & Galloway	145,800	147,830	-1.4	147,780
Dundee City	142,700	146,990	-3.0	145,460
East Ayrshire	120,630	120,010	0.5	120,310
East Dunbartonshire	110,760	108,510	2.0	108,250
East Lothian	91,280	89,850	1.6	90,180
East Renfrewshire	89,790	89,030	0.8	89,410
Edinburgh, City of	453,430	447,900	1.2	449,020
Eilean Siar ²	27,180	26,730	1.7	26,450
Falkirk	144,320	144,770	-0.3	145,270
Fife	350,400	348,330	0.6	349,770
Glasgow City	609,370	573,180	5.9	578,710
Highland	208,600	209,050	-0.2	208,920
Inverclyde	84,600	84,610	0.0	84,150
Midlothian	82,200	81,080	1.4	80,950
Moray	84,950	86,890	-2.3	87,000
North Ayrshire	138,850	136,110	2.0	135,820
North Lanarkshire	327,620	320,860	2.1	321,180
Orkney Islands	19,480	19,290	1.0	19,220
Perth & Kinross	133,620	135,150	-1.1	134,950
Renfrewshire	176,970	173,230	2.1	172,850
Scottish Borders	106,900	106,190	0.7	106,950
Shetland Islands	22,440	22,140	1.3	21,960
South Ayrshire	113,920	112,080	1.6	112,160
South Lanarkshire	307,400	302,050	1.7	302,340
Stirling	85,220	86,350	-1.3	86,200
West Dunbartonshire	94,600	93,840	0.8	93,320
West Lothian	156,690	157,300	-0.4	159,030

¹ Published in April 2001² Formerly Western Isles

Table 3 Numbers added in 1991 (aged 10 years) compared to difference between previously published and provisionally revised mid-2000 population estimates

Persons				Males				Females			
Age in 2001	Added in 2001	Added in 1991	Published 2000 MYE - Revised 2000 MYE	Age in 2001	Added in 2001	Added in 1991	Published 2000 MYE - Revised 2000 MYE	Age in 2001	Added in 2001	Added in 1991	Published 2000 MYE - Revised 2000 MYE
Total	198,987	159,095	56,400	Total	107,529	100,460	56,587	Total	91,458	58,635	-187
0	4,107	0	0	0	2,176	0	0	0	1,931	0	0
1-4	12,768	0	5,968	1-4	6,912	0	2,748	1-4	5,856	0	3,220
5-9	15,447	0	10,002	5-9	8,106	0	5,028	5-9	7,341	0	4,974
10-14	13,977	9,444	3,484	10-14	7,154	5,038	1,700	10-14	6,823	4,406	1,784
15-19	15,692	5,216	4,103	15-19	8,211	2,709	2,865	15-19	7,481	2,507	1,238
20-24	27,222	2,290	7,646	20-24	15,491	1,180	7,076	20-24	11,731	1,110	570
25-29	21,512	15,840	14,591	25-29	12,235	8,933	15,340	25-29	9,277	6,907	-749
30-34	19,671	39,344	18,472	30-34	11,557	26,166	16,271	30-34	8,114	13,178	2,201
35-39	15,529	32,692	14,223	35-39	8,641	24,475	12,875	35-39	6,888	8,217	1,348
40-44	12,477	17,292	7,069	40-44	6,458	13,696	7,090	40-44	6,019	3,596	-21
45-49	7,887	6,750	-1,791	45-49	4,852	4,760	-841	45-49	3,035	1,990	-950
50-54	7,885	5,184	-3,285	50-54	4,818	3,203	-2,555	50-54	3,067	1,981	-730
55-59	5,496	3,083	-5,167	55-59	3,252	1,527	-3,484	55-59	2,244	1,556	-1,683
60-64	4,910	3,083	-4,261	60-64	2,781	1,527	-2,919	60-64	2,129	1,556	-1,342
65-69	3,717	3,083	-5,121	65-69	1,414	1,527	-2,273	65-69	2,303	1,556	-2,848
70-74	3,580	4,046	-1,666	70-74	1,277	1,527	-701	70-74	2,303	2,519	-965
75-79	2,706	3,916	-2,053	75-79	1,045	1,397	-574	75-79	1,661	2,519	-1,479
80-84	2,584	3,916	-1,480	80-84	737	1,397	-629	80-84	1,847	2,519	-851
85+	1,820	3,916	-4,334	85+	412	1,397	-430	85+	1,408	2,519	-3,904